WAGE DIFFERENTIAL BETWEEN METROPOLITAN AND NON-METROPOLITAN AREAS: CASE OF UKRAINE

by

Inna Bisovetska

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Arts in Economics

National University "Kyiv-Mohyla Academy"
Economics Educational and Research Consortium
Master's Program in Economics

2006

Approved by _____
Mr. Serhiy Korablin (Head of the State Examination Committee)

_____

_____

_____

Program to Offer Degree Master's Program in Economics, NAUKMA

Date _____

National University "Kyiv-Mohyla

University"

Abstract

WAGE DIFFERENTIAL BETWEEN METROPOLITAN AND NON-METROPOLITAN AREAS: CASE OF UKRAINE

by Inna Bisovetska

Head of the State Examination Committee: Mr. Serhiy Korablin,
Economist, National Bank of Ukraine

The paper explores an urban wage premium and factors that cause it. In our research we tested three theories that should determine the size of urban wage premium and factors that cause it: omitted ability variable; higher productivity of urban worker and higher cost of living in big cities. On the basis of the data from Ukrainian Longitudinal Monitoring Survey 2003 we performed OLS estimation for ordinary wage regression, instrumental variable technique, and fixed effect regression. We found that the argument in favor of positive urban wage differential is controversial and urban wage premium cannot be fully explained by any of the hypothesis. There is evidence that returns on some human capital characteristics are higher in urban settlements; omitted ability revealed to be not important in determination of wage differentials; and migration in both directions: from rural to urban settlements and vice versa, does not influence on the salary within this sample.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my thesis advisor, Prof. Yuri Yevdokimov, for his feedback and valuable advice. Special thanks go to Prof. Tom Coupe, from whom I received many helpful comments on my research. I am also grateful for Olga Kupets and Volodymyr Vakhitov for their help in providing me with data. Finally, I would like to thank my colleagues Klos Nadiya and Viktor Flinta for their enormous patience.

Chapter 1

INTRODUCTION

Under assumptions of perfect competitiveness of both labor markets in metropolitan and non-metropolitan areas (when equal conditions and opportunities for workers and employers are established), and perfect labor mobility, the marginal product of labor seems to be higher in metropolitan areas than in non-metropolitan because in big cities the compensation in labor and housing markets tends to be higher than in the smaller ones. This fact attracts people and gives them an incentive to move from one region to another as well as from small towns to big cities in order to increase their well-being. And despite the fact that in the reality assumption of perfect labor mobility never holds (which prevents all rural population from instantaneous moving to the cities), still the urban population in metropolitan areas has steady tendency to grow. This is true for developed countries, where market economies operate smoothly and generate such compensation differentials in labor and housing markets for specific local amenities, the same time this remains an open question for countries in transition.

The purpose of our paper is to investigate whether there is a significant difference in the size of wage and housing compensation between metropolitan and non-metropolitan areas across Ukraine and to determine what factors cause wages to be higher in the large cities than in the small ones.

A good framework of studying wage differential was already developed and there is a great variety of papers where this topic was examined. Three the most

popular and widely applied theories which might explain the existence of compensation differential are summarized in *Consumer City* paper by Glaeser, Kolko and Saiz (2000), and we will test them to see whether they can explain wage difference between Ukrainian big cities and small towns.

The first one is that the urban wage premium is illusory and represents omitted ability variable. This hypothesis undermines that cities are somehow better than rural areas. Maybe cities are so attractive because of non-work related factors such as greater ability to social relations, or presence of more commodities that are particularly desired by individuals. Moreover, some local amenities may be attractive to such extent, that people will be still willing to reallocate even if their new place of settlement provides lower earnings. If it is so, then workers who live in cities for non-wage reasons and workers that have just moved to the metropolitan areas will not experience benefits of urban wage premium and for some of them wages may be even lower than before reallocation, while workers that have moved to the city in order to increase their well-being, would receive urban wage premium.

The second hypothesis is that the urban wage premium comes from a level effect where cities (either because of reduced transportation costs or because of urban externalities) enhance worker productivity. For example, cities provide more facilities to get education in schools, institutes, universities, etc. and facilities to acquire specific knowledge and to develop skills in a narrow range of activities. Cities accumulate human capital and each citizen has an easy access to it. So on condition that hypothesis holds, cities will not affect productivity of new workers immediately after the change of settlement. Only over time well educated workers or workers who acquired particular skills would receive an urban wage premium.

The last hypothesis is that the urban wage premium is the result of economy of scale in production in cities. Metropolitan areas can be considered as huge

markets and their density reduces transportation costs, also there is always easier access to transportation networks comparing to rural areas. Lower cost of moving inputs (outputs) allows firms to save money and increases overall productivity. On the other hand, higher wages and better market conditions attracts more people, but the land is limited and this drives up urban rents, therefore in cities wage premium is offset by higher cost of living. If it is really so, then all workers should receive higher wages benefit immediately after their settlement.

One important notion is that Ukraine is a transition country, so while doing the research we will check whether it is true that process of wage determination in transition economies has its own peculiarities. For example, we expect that compensation size (especially for state owned enterprises) may be determined not purely by market conditions but also by government policies, which is a very characteristic feature of early stage of transition and which also may lower compensation difference between urban and rural areas (Boeri, Terrel (2002). We may also find returns to education and experience to differ from those described by the authors of *Consumer City* paper. Because education and experience acquired under communism might not be perfectly applicable to market economy, therefore it is logical to expect higher wages for younger and well educated people, but not for older professionals (Nesterova and Sabirianova (1998)).

The remainder of the paper is organized as follows. In Chapter 2 we provide a theoretical framework for the concept of wage differential. And provide an overview of several fundamental studies that contributed to the formation of the theories which we use to explain wage differential. In Chapter 3 we introduce the econometric model used to estimate the size of wage differential and to see what influence it. In Chapter 4 we discuss data, estimation techniques and present

major estimates. Finally, we summarize the main findings and draw some conclusions in Chapter 5.

Chapter 2

LITERATURE REVIEW

In this chapter we provide a theoretical framework on the link between size of worker's compensation, his productivity and the place of residence and briefly describe some empirical researches that were done before.

A huge amount of literature studies wage differentials, the reasons which causes them and their consequences. However many papers focus on developed countries. For example, Glaeser, Kolko and Saiz (2000), investigate differences in the size of compensation in metropolitan and non-metropolitan areas. They give three possible explanations of the urban wage premium in United States: (1) big cities attract people for non-work related reasons which might have ambiguous effect on wage; (2) urban wage premium comes from the fact that cities enhance accumulation of knowledge and skills of its citizens and rises their productivity; (3) higher urban wages is a result of higher cost of living and is a compensation for negative externalities. Empirical evidence showed that omitted ability variable explains very little in urban wage premium. Testing between the remaining two possible explanations showed that the higher urban compensation to a larger extent is due to accumulation of experience.

An interesting research was made by Coelho and Shepherd (1979) where differences in regional prices and wages and the relationship between the cost of living and city size, and the determinants of regional industrial growth were examined for the United States in 1980. The authors constructed price indexes on the basis of prices for goods and services – composites of the consumption

bundle, wage compensations for particular occupations and household expenditures in 80 different cities and found out that there is strong positive relationship between wages and size of a city, and between size of a city and the cost of living, but cost of living has lesser impact on wage premium.

Conclusion made by Coelho and Shepherd is the good evidence in favour of the second hypothesis which explains urban wage premium by the higher productivity of metropolitan citizens.

Another research made by Ruhm (1990) who studied such productivity characteristics of individuals as experience and seniority and their returns on wages. Ruhm found out that both experience and tenure are crucial determinants of size of compensation. But the author also argued that cross sectional wage regression might overestimate the extent to which earnings increase with job seniority because they do not take into account that sometimes high-wage workers have lower level of mobility. He looked at the inter-firm earnings of employees and the frequency of job switching among workers in order to determine whether seniority premium results from job specific attributes or from the fact that the superior workers simultaneously receive higher wages and have lower rates of job switching. As a result he came to the conclusion that earnings increase not only due to seniority. This finding is especially important for professionals, because a person with longer tenure appears more vulnerable to lose the seniority premium in the case of changing the job.

In our research we would like to test whether such returns for tenure and experience are the same across all the cities of the country or whether they vary with the size of the city and population density.

Roback (1982) focused her study on the role of wages and rents in allocating workers to locations with different amounts of amenities. She came to the conclusion that if amenity is also productive than it drives up rent costs as well as wages, and regional wage differentials can be explained largely by these local attributes. Since a lot of firms are located in metropolitan areas and it can be considered as a productive amenity, this research supports the theory of higher overall productivity of big cities which allows paying higher wages to urban workers.

A lot of researches were dedicated to the impact of local amenities on compensation size on the decision of people to reallocate. An example is research made by Glaeser and Mare (1994). Authors found out that despite the widely spread point of view that urban wage premium is the main reason that attracts people to agglomerations, it is not always true. Empirically they proved that very often rents in metropolitan areas grow up faster than wages, suggesting that the demand for living in cities rises for reasons beyond rising wages. The authors distinguished between four particularly critical urban amenities which generate an increase in number of city inhabitants. First and the most obvious one is the presence of rich variety of consumer goods, services (like theatres, restaurants, etc.) and social partners (friends, potential spouses) which are strictly local and cannot be transported somewhere else. The second amenity is favourable climate conditions. The next one is the presence of such public services as schools, institutes, universities, police, etc. which is linked with urban growth rate. The last one is speed and ease with which individual in metropolitan are can move around the great variety of services and jobs available. All amenities mentioned above are determinants of quality of life, so the conclusion is that the high quality of life alone can cause an increase in urban population.

The paper described above supports our first hypothesis that omitted ability variable attracts people to big cities which usually provide huge variety of amenities. So we would like to check whether it holds across the territory of Ukraine.

Similar topic was developed by Blomquist, Berger and Hoehn (1988). They examined regional differences in amenities. The authors determined that climatic, environmental and urban conditions are key variables for determining indexes for quality of life in different locations and showed that quality of life is the most important factor considered in location decision made by people. So higher wages again seem to be not the only reason why people live in urban areas.

Special attention should be devoted to the transition countries, because transition economy has its own peculiarities which need to be considered while doing a research. Theories which hold for developed economies may not work in the same way in transition economies. For example Boeri and Terrel (2000) investigated how transition affects wages and allocation of labour force in post-Soviet countries. Their study was focused on institutional determinants of GDP, employment and labour reallocation which are reasons for significant variations in the transition paths of post soviet countries. And authors empirically proved that in different countries during transition from centrally-planned to market economy same theory works in a different way for every special country.

At the same time some factors may have similar effect on wage differentials in both, developed and transition countries.

Berger, Blomquist and Sabirianova (2003) examined the existence of compensational differentials for location specific amenities in Russian cities. Empirical evidence suggests that workers are compensated for differences in

climate, environmental conditions, high crime rates and ethnic conflicts even after controlling for worker characteristics, differences in market conditions and "regional coefficients" paid by Russian government to compensate workers for living in regions with the lowest quality of life. Also quality of life index, constructed by the authors, is positively correlated with net migration into a region, suggesting workers are attracted to amenity-rich locations. So transition economies may be able to generate wage difference between places with different quantity of amenities as well as the already developed ones.

Jurajda (2004) focused his research on returns on education in transition economies. He tested for the presence of human capital spillovers and regional variations in returns to education in Czech Republic. He found out that the local concentration of human capital in the economy that inherited dramatic variations in local skills endowments from the planned economy and wages by level of education and lack returns to education. Empirically he showed that the difference in the level of development in regions and regional wages is not due to the local concentration of skills but due to the presence of availability of local educational opportunities. According to his findings the level of education has positive impact on size of the wage as well as the presence of opportunities to get this education. Since metropolitan areas are usually scientific and educational centers, one may expect higher wages in big cities.

Very often variables that are significant determinants of wage differentials in developed countries are less significant in transition ones. The bright example of such determinant may be human capital.

The topic of investment in human capital is discussed by Herasym (2004). The author showed that the theory which states the following: the better education and experience – the greater payoff is expected in future, always works in the

developed countries, but not in transition economies. On the basis of Ukrainian Longitudinal Survey they showed that in Ukraine the winners of transition became younger people, but not the older and more experienced ones and newly established private firms tend to value education more than experience. This is probably because education and experience acquired during soviet time may not be perfectly applicable to new conditions of market economy.

Also some problems may appear in investigation wage differentials across Ukrainian cities. Gryshyna (2001) investigates what factors cause wage areas and influence on its length across Ukraine. Her analysis suggests that the most powerful determinants of wage areas are regional location, industrial affiliation and type of ownership. And since on early stages of transition (1990s) majority of enterprises were state-owned, size of wages was set by the government and did not vary much across the territory of the country, so we may expect not very significant size of wage differentials.

Chapter 3

METHODOLOGY

I our study we will the model, very similar to the one, described by Glaeser, Kolko and Saiz in the paper "*Consumer City*":

$$\ln W_{it} = \alpha_i + \beta_1 C_{it} + \beta_2 T_{it} + \beta_3 X_{it} + \beta_4 X_{it} C_{it} + \beta_5 X_{it} S_{it} + \beta_6 X^2_{it} + U_{it}$$

α_i – is a constant term which captures an individual specific productivity effect (individual ability) as well as not perfect mobility of Ukrainian labor market.

$\ln W_{it}$ - is natural logarithm of the nominal wage of the worker.

C_{it}- is a dummy variable that takes value of 1 if individual lives in a city with high density of population (500 000 inhabitants and above).

β_1 – indicates how earnings of worker changes only because he lives in a city with high density of population, in comparison to the earnings of rural worker.

T_{it}- is another dummy variable for urban types of settlements. It takes value of 1 if an individual lives in a town (from 20 000 to 499 000 inhabitants).

β_2 – indicates how living in a town changes earnings of the worker again, in comparison to the earnings of rural worker.

X_{it} - indexes the set of observable human capital characteristics that are likely to influence individual's wage (i.e. years of schooling, experience, tenure and other human capital components).

β_3 – represents the price of those human capital characteristics in Ukrainian labor market. It is the same for all individuals in the sample independently of their place of satlement.

$X_{it}C_{it}$ and $X_{it}T_{it}$ reflect interaction effect between human capital characteristics and individual's urban residence.

β_4 – reflects the wage premium for certain human capital characteristics which is paid for workers in big cities.

β_5 – reflects the wage premium for the same human capital characteristics which is paid for workers in small cities and in towns.

X^2_{it}– are introduced into regression in order to assure the concavity of experience and age earning profiles that is generally suggested by human capital theory and extensively justified by empirics. We expect the coefficients for age^2 and $exper^2$ to be negative.

U_{it} - is a statistical residual that captures unobserved innate and other characteristics that influence worker's salary but are not included into equation. We assume that U_{it} is normally distributed and has constant variance.

For hypothesis testing we use ordinary least squares technique, namely: for wages we run ordinary wage regression described above. We focus our attention on

estimates for C and estimates for T dummy variables which give us the size of urban wage premium.

In order to see on returns to human capital characteristics in metropolitan areas and in rural areas we add individual characteristics such as years of schooling and experience. Both, education and experience are independent from urban labor market, so it makes possible to see whether in metropolitan areas returns to experience and education are really higher than in non-metropolitan areas. We expect the sign of coefficients to be positive because according to the theory, wages rises with experience and education. However we allow that these individual characteristics might have lesser impact on wage size in Ukraine than in the case of developed countries.

Next step, we add job characteristic: tenure and numbers of trainings that individuals have undertaken. Again, we expect positive signs of coefficients for both urban and rural workers and are particularly interested whether returns on tenure and trainings are higher in metropolitan than in non-metropolitan areas. An important variable, which we also include, is average occupational education. It can be considered as job characteristic and at the same time it will control for background characteristics of individual in one occupational group.

Another hypothesis that we have to test is omitted ability hypothesis. In this case OLS technique cannot be applied because we may expect C and T to be correlated with error term, therefore OLS regression will give us biased and inconsistent coefficients. We try to solve this problem by estimation of urban wage premium with the help of instrumental variable that is unrelated to individual ability but that do increase likelihood that individual will reside in big city. As instrument we use place of birth of individual which is strictly exogenous (unlike place of current residence which is endogenous).

All steps described above might help to determine whether it is true that omitted ability can really drive wages up, but they are not be sufficient to distinguish between the other two hypotheses. In order to determine what exactly cause urban wage premium: higher individual or overall productivity in metropolitan areas need to see how fast wages of migrants increase over time. That is why we observe wages of dwellers for some period of time before and after their last change in residence. Also we introduce into regression two additional dummy variables for individuals who moved from metropolitan to non-metropolitan areas and from non-metropolitan to metropolitan areas, because migration in both directions may cause changes in wages.

Chapter 4

DATA DESCRIPTION

Data for our analysis come from Ukrainian Longitudinal Monitoring Survey in 2003[1] that contains information on the representative sample of 8641 individuals drawn out of the working age population between 15 and 72 years old. The individual ULMS questionnaire provides detailed information on individual's employment, size of compensation, level of education, age, skills, time allocation, etc and, therefore, is suitable for our research.

The data used in the research is restricted to sample consisting 3260 individuals that were employed during 2003 are reported their earnings in the section E "Main Job and Second Jobs in the Reference Week". We also make use of the retrospective part of ULMS questionnaire. Section C "Main Jobs in 1896, 1991, 1997, 1998-2003" provides us with the information on the changes in income of individuals. From the Section H "Changes of Residence" we obtained information about all changes of settlements made by individuals from our sample.

For the purpose of our research, namely for the possibility to compare the size of worker's compensation across different types of inhabited localities, we divided all individuals in the sample into three categories. The first category consists of individuals who live in big cities with high density of population (more than

[1] ULMS 2003 was carried out by the Kiev International Institute of Sociology at the request of IZA, Centre for Economic Reform and Transformation (CERT), Economics Education and Research Consortium (EERC)-Ukraine, Leuven Institute for Transition Economics (LICOS), Rheinland –Westfaelisches Institute feur Wirtschaftsforshung (RWI)-Essen, and the William Davidson Institute (WDI).

500 000 people); the second category includes individuals from medium and small cities (from 20 000 up to 499 000 inhabitants); and villages (less than 20 000 inhabitants). For all geographical categories we used only information of where individual lived, not where he worked.

Table 1 gives main characteristics of the variables which are used in the analysis, namely means, standard deviations, minimum and maximum values.

We have 22 820 observations (3260 individuals are observed during 1986-2003) in our sample. Urban population is represented but two categories: City and Town and rural population is represented by the Village category. Of these 22 820 observations, 5 559 observations are City observations, 7 760 are Town observations, and, finally, 6 895 observations are Village observations.

As a dependent variable we used salary. Since our analysis considers wages of individuals over the period of seven years, an increase in earning might be explained to the larger extent by inflation, therefore taking a logarithm is a necessary step. In order to deflate and to eliminate the effect of inflation we uses log of nominal wages and also we add time dummies. This should eliminate the effect of inflation, economic shocks or unrelated business cycle effects.

Before performing any empirical analysis we can see that there is a difference in log of salary between different types of inhabited localities, and it increases with the size of the settlement. The difference in log of salary between Town and Village is 0.194 and the difference between City and Village is 0.253. Between City and Town the difference in log of salary is very small 0.059.

Explanatory variables we used in our analysis describe individual characteristics of workers and are widely described in the theory as main determinants of salary.

The first one is experience. We calculated actual amount of years during which a person was employed. The average individual in our sample has 13 years of experience, but there is a little difference in this variable across geographical locations. The longest experience has a person from Town category –13.7 years. this is because experience is usually positively correlated with age. And as we can see from the table 1, workers in Town are slightly older than workers in other geographical categories. However, age also does not significantly vary across categories and deviations from the average age in the sample (13.5) are minor.

Years of schooling differ across categories. Average person in the sample has 12.9 years of education. There is no difference in education between urban categories. In City average citizen has 13.24 years of schooling, in town – 13.21 years. However in the Village average education is 12, 57 and is less than in urban settlements.

Tenure is a variable that describes number of years that individual have worked on the current job. It is a special type of human capital, because it is especially valuable for the current employer. Tenure is correlated with experience and age, but it is logical, because experience increases with age and experience includes tenure. From the sample statistics we see that average duration of tenure in the sample is 7.4 years. In Town tenure is the longest one – 7.7 years, while in Village category tenure of employees does not go far from the average value. However in the City category, tenure is shorter than in the other categories - only 6.7 years. This means that in cities with high density of population people tend to change jobs a little bit more than in towns and villages.

Training variable gives the number of courses and trainings that individual have undertaken. Under trainings in ULMS 2003 were undermined foreign language courses, upgrading skills courses, use of computer, driving, administration of firm courses, courses for salesmen, etc. We assume that any kind of training has

positive impact on worker's earnings, since it increases human capital and, therefore, is of value for the employer. The number of trainings varies in the sample from 0 to 3.

Some unavoidable problems may appear, for example, reference bias. Our research is built on data from the retrospective part of ULMS and there is always possibility that people did not remember exactly therefore the numbers and dates which they reported are incorrect. Another problem is that data is unbalanced. Many respondents did not report their wages in 1986-1997 year and this reduces the number of my observations. Also in Section H "Changes of Residence" people did not report the name of the settlement where they were born or where they had moved to. We need this data in order to create dummy variables for different types of settlements and in order to distinguish the direction of migration. Absence of this data reduces number of observation even more. We assume that omitted observations are distributed randomly; therefore unbalanced panel data problem is not severe.

Chapter 5

EMPIRICAL ESTIMATION

Since the primary interest of our research is urban wage differential, we start our analysis with estimation by OLS. This allows us to see what factors influence on individual's salary. Also we add dummy variables for urban type of settlements in order to see whether living in a city or a town affects salary too. Introducing human capital characteristics into the regression and their interaction with urban dummy variables allows us to test one of our hypotheses about higher returns on human capital characteristics in metropolitan areas.

The obtained results revealed that city wage differential coefficient is statistically significant even at 1% confidence interval and equals to 0.53. Since we have log-lin model and our dependent variable is log of nominal wages, we have to interpret estimated coefficient not in numbers, but as a percentages. So workers, who live in the cities with high density of population, receive by 53% more than workers in villages or in very small towns.

It is necessary to mention here if wage differential is statistically significant and positive this should not lead to the conclusion that living in the city is more attractive that in village. In our analysis we do not control for the expenditures of individual on housing and consumption, therefore we cannot claim that higher salary increases individual's wealth. If cost of living in the cities significantly greater than in the villages, individuals in the cities may be in a worse conditions even with higher salaries that village inhabitants. The aim of our analysis is only to

see whether wages of urban workers are higher and find appropriate explanation for this phenomenon, but not to compare well-being of urban and rural workers.

Coefficient near Town dummy is insignificant; therefore we can say that there is no significant difference in earnings of workers from middle-sized cities and villages.

We see that coefficients near age, education, tenure and training variables are statistically significant and have positive sign. So these individual characteristics increase earnings in Ukrainian labour market. However they affect workers across all types of settlements in the same way. In order to see the returns on human capital characteristics in urban labour market we have to focus our attention on interaction effects of City and Town dummies with vector of human capital characteristics.

Results of the regression revealed that in City returns to human capital characteristics are not as high as we expected. Moreover, it tuned out than in Town returns on education are higher than in city. Each additional year of schooling adds 0.014% to the salary if a person lives in a city and 0.041% to the salary if he lives in a town. So if, for example , individual with 10 years of education receive by 1,4% is he lives in the city and by 4,1% more if he lives in the town, comparing to the person from the village with 0 years of schooling.

Coefficient for experience has negative sign and is statistically significant. However, if we look on returns on experience in cities, it will be positive, but it will increase wages by small amount: by .01% with each additional year of experience. This means that 10 year of experience will raise salary of a worker from city only by 1%, comparing to the person from the village with 0 experience. Living in a town will not bring any extra payments for experience

comparing to the wages in village. Such a small impact of experience on individual earnings force us to think, that really experience, acquired under centrally planned economy is not perfectly applicable for the conditions of market economy and therefore are of little value for employers.

Returns on tenure are very small for City (less than 0.01%) and are insignificant for Town category. Trainings coefficients are significant at 5% confidence interval for both urban categories of settlements; however again, their effect on wages is minor. Living in a city will bring 0.08% to the salary with each next training and living in a town will bring 0.05% to the salary with each additional training.

Coefficient of age and gender dummy are statistically significant. Returns on age and on being man do not differ between Town and City, but in the both categories are higher than in village.

When we are dealing with omitted ability hypothesis, we may expect that City and Town variables are not strictly exogenous and are correlated with error term. IV estimation is the most frequently used method to handle the problem of endogeneity and possible measurement error. In our case we need to find such an IV, which will be correlated with urban settlement dummy and not correlated with error term, i.e. it should increases likelihood of living in urban settlement, but is not correlated with error term. As an instrument we use place of birth of individual, which is strictly exogenous.

The results of the regression with IV are not directly comparable t the results of previous one because dummy variable for urban types of settlement is different. Now we use City_new dummy variable, which takes value of 1 if individual, lives in a City or Town. Variables City_new_educ and City_new_exper which stands

for interaction effect between urban dummy variable and vector of human capital characteristics (education and experience respectively) should be also instrumented, since they can correlate with error term too.

From the result of the regression we see that urban wage premium is statistically significant at 1% confidence interval, and has positive sign. This means that worker in a city or town receives salary by 290% more, that the worker with the same characteristics in a village. So urban employee receives salary almost 3 times higher than rural employee.

Another application for the obtained results is that omitted ability hypothesis does not alone explain urban wage differential, since its coefficient remains significant eve after using IV.

As for the OLS estimation, the picture is not very promising too.
Breusch-Pagan test for heteroskedasticity showed, that there is such a problem in our sample. Since p value is equal to one, we reject the null hypothesis that variances in the model are constant. Also Ramsey test provide evidence that there is omitted variable in regression. Again p value is equal to one; we reject the null hypothesis that there are no omitted variables in our model. Taking into account results of the tests we cannot rely on obtained earlier estimates since they are biased and inconsistent. In order to solve these problems we will use panel data. It will help to reduce omitted variable bias and will reduce heterogeneity across individuals.

However, another problem is that coefficients estimated by pooled OLS are also not reliable because the error term may consist of two parts: fixed effects that vary only over individuals and random effects that vary both over time and over individuals. Even if correlation between fixed effect and all explanatory variables

is equal to 0, we still have a problem of serial correlation and our standard errors are biased. So before running any regression on our econometric model we have to make several tests first.

In order to see what estimation technique is better OLS or fixed-effects we perform Breusch-Pagan test. The null hypothesis which is tested is there are no variations in error term and estimations obtained by pooled OLS are unbiased. From the results we see that p value is very low (0,000), therefore we have to reject the null hypothesis and use random effect model.

While running fixed effect regression that correlation between error term and explanatory variables is indeed present: corr(u_i, Xb) = -0.8875, and all coefficients are jointly equal to 0 (F test is very high) so using fixed effect regression seems to be reasonable. However with fixed effect estimation education and average education within one occupational group, dummy variable for man dummies for 1998, 1999, 2002, 2003 years and are dropped from the regression due to multicollinearity. Also since only one variable - Salary in our analysis really varies over time, other variables such as experience, tenure, education, etc. become redundant because they change over time in a similar way for all individuals in the sample(i.e. $expr_{it+1}$ - $expr_{it}$=1). Also urban wage premium becomes insignificant comparing to the random effect estimates.

To see what regression is better (RE vs FE) we need to prepare Hausman test. Under the assumption of no correlation we should obtain same results for both FE and RE, however we see that standard errors are very small (statistically insignificant), so we cannot reject the null hypothesis that there is no correlation, therefore RE will give us biased estimates and we stop our choice on FE estimation.

FE estimation does not provide us with sufficient explanation of wage differential. On the contrary, now coefficients near City and Town dummy variables are statistically insignificant. However returns on some human capital characteristics are significant. For trainings returns are highest in the town –here person receives extra 4,5% of salary for each additional training comparing to the worker in the village.

So as we have seen, returns on human capital is higher in the urban labor market, however at this stage we can not say exactly whether it is only human capital drives up wages in cities and towns or maybe higher cost of living determine urban wage premium as well. In order to distinguish between these two hypothesizes, we have to observe how wages of employees change over time. Also we should separate from our sample individuals, who have changed piece of residence and to see whether they have changed types of settlement during their migration. We created two dummy variables migr_vc and migr_cv. Migr_vc takes value of 1 if individuals moved from village to city (city now stands for both, towns and cities) at the year when change of residence took place. Migr_cv dummy variable reflects migration in the opposite direction.

We focus our attention on coefficients near migr_vc and migr_c, which should tell us how change in settlement influence on wages of dwellers within the year after migration. As we see form the regression results, both coefficients are insignificant. In order to observe migrants some time before and after migration we introduced lags and leads into regression.

As a regression results revealed, the only significant coefficient at 5% confidence interval is coefficient near L2migr_cv variable. Interpretation is the following: after 2 years since individual migrates from city to the village he receives 27% increase in salary. This is probably because such person was not very successful in the city and therefore he /she moved to the village, where there are more chances

for him/her to earn more. Since we do not observe increase in the salary of dwellers immediately after migration, urban wage premium cannot be explained by higher costs of living in cities and towns, comparing to villages.

Chapter 6

CONCLUSION

The main idea of our research was to investigate whether there is a significant difference in the size of wage and housing compensation between metropolitan and non-metropolitan areas across Ukraine and to determine what factors cause wages to be higher in the large cities comparing to small towns and villages. For this purpose we used data on individual level from the Ukrainian Longitudinal Monitoring Survey that was conducted in 2003. The ULMS contains detailed information about individual characteristics of the respondents, their income and changes in the place of residence which is especially important for our research. retrospective part of ULMS allows us to observe individuals over a period of time since 1986 up to 2003.

In our research we tested three theories that should help us to determine the size of urban wage premium and find what factors cause it. The first hypothesis assumes that urban wage premium is illusory and represents omitted ability variable. It undermines that cities are somehow better than villages because of the strictly local amenities that they provide (i.e. commodities that are particularly desired by consumers, entertaining, communication, community, etc.). The second hypothesis tells us that in big cities workers are more productive, than in small towns or villages and returns to human capital characteristics are higher in urban areas. The last one suggests that urban wage premium results from higher cost of living in cities with high density of population.

For hypothesis testing we use OLS estimation for ordinary wage regression, but obtained results were not reliable because of heteroscedasticity and omitted

variable problems. Also theory suggests that there is endogeneity problem, therefore we applied instrumental variable technique. In order to solve the problem of omitted variable and to observe changes in earnings of individuals over time, we used panel data and according to Hausman specification test results we run fixed effects regressions.

We found that the argument in favor of positive urban wage differential is controversial; however it appeared to be significant in majority of regressions. We also found that urban wage premium cannot be fully explained by any of the hypothesis mentioned above. There is an evidence that returns on some human capital characteristics (education, training) are higher in urban settlements, but not to such an extent to explain urban wage premium. Omitted ability revealed to be not important in determination of wage differentials. Another important finding is that migration in both directions: from rural to urban settlements and vice versa, does not influence on the salary within this sample.

There is a positive impact of migration on wages only in the case when person moves from city to village. Two years after the change of place of residence, the salary of individual increases by 27%. This is probably because people who are not successful in cities tend to move to settlements with lover density of population, where he can become more effective worker.

One possible explanation why obtained results does not fully explain urban wage differential is violation of crucial assumptions of the used model. Fixed effect estimation provides efficient estimates on condition that there is no autocorrelation and heteroscedasticity in the sample, however we cannot be sure that there is no such problem in our case. Unfortunately Stata can not perform test for autocorrelation and heteroscedasticity for our sample because of the huge number of observations and run regression with correction for autocorrelation

and heteroscedasticity. So only because of technical problems a better analysis of the topic can not be performed.

BIBLIOGRAPHY

Glaeser, Edward; Kolko Jed; Saiz, Albert, *Consumer City* (July 2000). NBER Working Paper # 7790.

Philip R.P. Coelho;JamesF.Shepherd, *The Impact of Regional Differences in Prices and Wages on Economic Growth: the United States in 1980*, Journal of Economic History, Vol. 39, #1 The Tasks of Economic History (Mar. 1979). 69-85

Glaeser Edward; Mare David, *Cities and Skills* (May 1994). NBER Working Paper # 4728.

Blomquist Glenn, Berger Mark, And Hoehn John, *New Estimates of Quality of Life in urban Areas, The American Economic Review,* Vol.78, #1, (Mar. 19880)

Ruhm, Christopher, *Do Earnings Increase with Job Seniority?*, *The Review of Economics and Statistics,* Vol. 72, #1 (Feb., 1990).

Roback Jenifer, *Wages, Rents and the Quality of Life, The Journal of Political Economy, Vol.90, #6 (Dec. 1982), 1257-1278*.

Tito Boeri; Katherine Terrel, *Institutional Determinants of Labor Reallocation in Transition, The Journal of Economic Perspectives,* Vol.16, # 1 (Winter 2002)

Herasym Galina *Private Investment in Human Capital:Effect on Earnings in Ukraine* EERC MA Thesis, (2004)

Jurajda, Stepan, *Are There Increasing Returns to Local Concentration of Skills? Evidence on Wages and Returns to Education in Transition,* CERGE-EI paper collection.

Mark C. Berger; Glenn C. Bloomquist; Klara P. Sabirianova, *Compensating Differentials In Emerging Labor And Housing Markets: Estimates of Quality of Life in Russian Cities.* IZA Discussion Paper No.900, (October 2003),

Olena Gryshyna The Incidence of Wage Areas in Ukraine: Evidence on Enterprise and Household Data. EERC MA Thesis,(2001)

APPENDIX A

TABLE 1. **SAMPLE STATISTICS**

		Total	City	Town	Village
Number of Observations		22820	5559	7760	6895
Log of Wages	Mean	5.338	5.463	5.404	5.21
	Std.Dev	0.63	0.66	0.59	0.62
	Min	0	0	0	0
	Max	11.156	11.156	8.517	8.29
Experience (exper)	Mean	13.194	12.62	13.681	13.11
	Std.Dev	11.825	12.17	12.05	11.55
	Min	0	0	0	0
	Max	60	60	60	57
Tenure (tenure)	Mean	7.367	6.683	7.666	7.503
	Std.Dev	10.116	10.07	10.121	10.104
	Min	0	0	0	0
	Max	57	54	53	57
Training (training)	Mean	0.084	.07	0.092	0.086
	Std.Dev	0.317	.028	0.33	0.325
	Min	0	0	0	0
	Max	3	3	3	3
Education (educ)	Mean	12.952	13.235	13.212	12.57
	Std.Dev	2.27	2.48	2.27	2.083
	Min	0	0	0	0
	Max	21	21	21	21
Average Education in Occupational Group (occup_educ)	Mean	12.952	12.956	12.952	12.949
	Std.Dev	0.108	0.105	0.107	0.11
	Min	12.41	12.41	12.41	12.41
	Max	13.14	13.14	13.14	13.14
Age	Mean	35.533	35.497	36.233	34.973
	Std.Dev	11.89	12.559	12.056	11.312
	Min	15	15	15	15
	Max	70	70	70	70

APPENDIX B

Regression Results

TABLE 2. **OLS ESTIMATION OF URBAN WAGE DIFFERENTIAL**

Source	SS	df	MS	Number of obs =	16661
Model	1376.7061	29	47.4726241	F(29, 16631) =	151.59
Residual	5208.25942	16631	.31316574	Prob > F =	0.0000
Total	6584.96552	16660	.395256034	R-squared =	0.2091
				Adj R-squared =	0.2077
				Root MSE =	.55961

lnsalary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
town	-.0228343	.0747885	-0.31	0.760	-.1694276	.1237591
city	.5280651	.0758036	6.97	0.000	.379482	.6766483
age	.030524	.0029055	10.51	0.000	.0248289	.036219
educ	.0366785	.0033598	10.92	0.000	.030093	.043264
exper	-.0043585	.0017164	-2.54	0.011	-.0077228	-.0009942
tenure	.0054369	.0010509	5.17	0.000	.0033771	.0074968
training	.1569873	.0190557	8.24	0.000	.119636	.1943385
male	.3999881	.0137026	29.19	0.000	.3731296	.4268465
age2	-.0004753	.0000377	-12.60	0.000	-.0005492	-.0004014
exper2	.0000352	.0000366	0.96	0.336	-.0000365	.0001069
occup_educ	.046874	.0402546	1.16	0.244	-.0320294	.1257773
city_educ	-.0228273	.0048772	-4.68	0.000	-.0323872	-.0132674
town_educ	.0040898	.0047942	0.85	0.394	-.0053074	.013487
city_exper	.0056461	.001888	2.99	0.003	.0019455	.0093467
town_exper	-.0006942	.0017823	-0.39	0.697	-.0041877	.0027993
city_tenure	-.0050411	.0016692	-3.02	0.003	-.0083129	-.0017694
town_tenure	.0024567	.0015409	1.59	0.111	-.0005637	.0054771
city_train~g	-.0767745	.0336968	-2.28	0.023	-.1428239	-.0107251
town_train~g	-.1103303	.0285363	-3.87	0.000	-.1662646	-.054396
city_age	.0005293	.0014771	0.36	0.720	-.002366	.0034245
town_age	.0041359	.001404	2.95	0.003	.0013839	.0068878
city_male	.0227427	.0223759	1.02	0.309	-.0211165	.0666018
town_male	.0239466	.020648	1.16	0.246	-.0165258	.0644189
y98	.0526535	.0184325	2.86	0.004	.0165239	.0887832
y99	.1527444	.0180712	8.45	0.000	.1173229	.1881659
y00	.2305768	.0177149	13.02	0.000	.1958538	.2652998
y01	.3160657	.0173792	18.19	0.000	.2820006	.3501307
y02	.389871	.0168904	23.08	0.000	.3567641	.4229779
y03	.3998713	.0166672	23.99	0.000	.3672017	.4325408
_cons	3.251302	.5226867	6.22	0.000	2.22678	4.275823

Omitted category :village, y97

APPENDIX C

Regression Results

TABLE 3. **INSTRUMENTAL VARIABLES (2SLS) REGRESSION**

```
      Source |       SS       df       MS              Number of obs =    7565
-------------+------------------------------           F( 18,  7546) =   71.49
       Model | -392.595205    18 -21.8108447           Prob > F      =  0.0000
    Residual |  3239.20335  7546  .429260979           R-squared     =       .
-------------+------------------------------           Adj R-squared =       .
       Total |  2846.60814  7564  .376336349           Root MSE      =  .65518

------------------------------------------------------------------------------
    lnsalary |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    city_new |    2.902291   .7587472     3.83   0.000     1.414935    4.389647
city_new_e~c |   -.1268713   .0433468    -2.93   0.003     -.211843   -.0418995
city_new_e~r |   -.0243977   .0095922    -2.54   0.011     -.043201   -.0055943
         age |    .0228908   .0056277     4.07   0.000     .0118589    .0339226
        educ |    .1076329   .0254286     4.23   0.000     .0577857    .1574801
       exper |      .00769   .0058358     1.32   0.188    -.0037498    .0191298
      tenure |    .0048464   .0011402     4.25   0.000     .0026114    .0070814
    training |    .1040827   .0218752     4.76   0.000     .0612012    .1469641
        male |    .4061008   .0172815    23.50   0.000     .3722242    .4399773
        age2 |   -.0004185   .0000721    -5.80   0.000    -.0005598   -.0002771
      exper2 |    .0001622   .0000641     2.53   0.011     .0000365    .0002879
  occup_educ |   -.1258784   .0859094    -1.47   0.143    -.2942849     .042528
         y98 |    .0585412    .031365     1.87   0.062    -.0029429    .1200253
         y99 |    .1782025   .0309707     5.75   0.000     .1174913    .2389136
         y00 |    .2710061   .0305535     8.87   0.000     .2111126    .3308995
         y01 |    .3532077    .030081    11.74   0.000     .2942407    .4121748
         y02 |    .4249179   .0292706    14.52   0.000     .3675393    .4822964
         y03 |    .4213857   .0289289    14.57   0.000      .364677    .4780944
       _cons |    4.173843   .9797054     4.26   0.000     2.253347    6.094338
------------------------------------------------------------------------------
Instrumented:  city_new city_new_educ city_new_exper
Instruments:   age educ exper tenure training male age2 exper2 occup_educ y98
               y99 y00 y01 y02 y03 born_type born_type_educ born_type_exper
------------------------------------------------------------------------------
Omitted category :village, y97
```

APPENDIX D

Regression Results

TABLE 3. **FIXED EFFECTS REGRESSION**

```
Fixed-effects (within) regression               Number of obs      =     16661
Group variable (i): id                          Number of groups   =      3260
R-sq:  within  = 0.2417                         Obs per group: min =         1
       between = 0.0014                                        avg =       5.1
       overall = 0.0061                                        max =         7
                                                F(27,13374)        =    157.88
corr(u_i, Xb)  = -0.8875                        Prob > F           =    0.0000
------------------------------------------------------------------------------
    lnsalary |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        town |  -.6112764   .5453562    -1.12   0.262    -1.680252    .4576989
        city |  -.8802316   1.198331    -0.73   0.463     -3.22913    1.468666
         age |  -.0299603   .0230454    -1.30   0.194    -.0751326    .0152119
        educ |  -.2019758   .3749634    -0.54   0.590    -.9369572    .5330055
       exper |  -.0044168   .0097683    -0.45   0.651     -.023564    .0147304
      tenure |  -.0093135   .0056522    -1.65   0.099    -.0203925    .0017656
    training |   .1353581   .0183079     7.39   0.000      .099472    .1712441
        male |  (dropped)
        age2 |  -.0007303   .0000881    -8.29   0.000     -.000903   -.0005575
      exper2 |  -.0001002   .0000834    -1.20   0.229    -.0002636    .0000632
   city_educ |   .0430654   .0851835     0.51   0.613    -.1239064    .2100371
   town_educ |   .0422584    .041401     1.02   0.307    -.0388934    .1234101
  city_exper |   .0017915    .016361     0.11   0.913    -.0302784    .0338614
  town_exper |   .0113074   .0101691     1.11   0.266    -.0086254    .0312402
 city_tenure |   .0130067   .0091427     1.42   0.155    -.0049142    .0309277
 town_tenure |   .0100677   .0072124     1.40   0.163    -.0040696    .0242051
city_train~g |  -.1633857   .0312395    -5.23   0.000    -.2246196   -.1021518
town_train~g |  -.0904588   .0277331    -3.26   0.001    -.1448196   -.0360979
    city_age |   .0113659   .0138849     0.82   0.413    -.0158505    .0385824
    town_age |  -.0082925   .0070817    -1.17   0.242    -.0221736    .0055886
   city_male |  -.1207685   .3076559    -0.39   0.695    -.7238176    .4822806
   town_male |   .3098835   .1530382     2.02   0.043     .0099071      .60986
         y98 |   .1514906   .0248054     6.11   0.000     .1028686    .2001126
         y99 |   .3301455   .0465842     7.09   0.000     .2388339    .4214571
         y00 |   .5011721   .0690991     7.25   0.000     .3657281     .636616
         y01 |   .6888694   .0916613     7.52   0.000     .5092004    .8685384
         y02 |   .8622775   .1143072     7.54   0.000     .6382191    1.086336
         y03 |   .9754079   .1370269     7.12   0.000     .7068157       1.244
       _cons |   9.718553   4.937484     1.97   0.049     .0403852    19.39672
-------------+----------------------------------------------------------------
     sigma_u |  1.2849713
     sigma_e |  .30418832
         rho |  .94693382   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Omitted category :village, y97
```

APPENDIX E

Regression Results

TABLE 3. **FIXED EFFECTS REGRESSION. ANALYSIS OF DWELLERS**

```
Fixed-effects (within) regression               Number of obs      =      6766
Group variable (i): id                          Number of groups   =      2615
R-sq:  within  = 0.1655                         Obs per group: min =         1
       between = 0.0023                                        avg =       2.6
       overall = 0.0035                                        max =         3
                                                F(20,4131)         =     40.96
corr(u_i, Xb)  = -0.8943                        Prob > F           =    0.0000
------------------------------------------------------------------------------
    lnsalary |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        town |   .1226549   .1979562     0.62   0.536    -.2654458    .5107556
        city |    .024037   .2296258     0.10   0.917    -.4261532    .4742273
         age |  -.0596056    .034865    -1.71   0.087    -.1279599    .0087486
        educ |  (dropped)
       exper |   .0221794   .0184612     1.20   0.230    -.0140146    .0583733
      tenure |  -.0359408   .0105612    -3.40   0.001    -.0566464   -.0152353
    training |   .0541283   .0215266     2.51   0.012     .0119245     .096332
        male |  (dropped)
        age2 |  -.0005872   .0002322    -2.53   0.011    -.0010424   -.0001319
      exper2 |   3.34e-06   .0002164     0.02   0.988     -.000421    .0004277
  occup_educ |  (dropped)
     migr_cv |   .1679679   .1410373     1.19   0.234    -.1085412     .444477
   F1migr_cv |   .0773044    .144064     0.54   0.592    -.2051386    .3597473
   F2migr_cv |  -.0359897   .1114577    -0.32   0.747    -.2545068    .1825274
   L1migr_cv |    .146641   .1218744     1.20   0.229    -.0922985    .3855805
   L2migr_cv |   .2682155   .1047342     2.56   0.010     .0628802    .4735509
     migr_vc |  -.1452654   .1570204    -0.93   0.355    -.4531099     .162579
   F1migr_vc |  -.0910742    .107878    -0.84   0.399    -.3025732    .1204247
   F2migr_vc |  -.1418935   .1059287    -1.34   0.180    -.3495709    .0657839
   L1migr_vc |  -.2087992   .1439144    -1.45   0.147     -.490949    .0733505
   L2migr_vc |  -.1778442   .1176085    -1.51   0.131    -.4084201    .0527318
         y98 |  (dropped)
         y99 |  (dropped)
         y00 |   .1953922   .0391838     4.99   0.000      .118571    .2722135
         y01 |   .4076249   .0773125     5.27   0.000     .2560507    .5591991
         y02 |  (dropped)
         y03 |  (dropped)
       _cons |   8.132956   1.283324     6.34   0.000     5.616951    10.64896
-------------+----------------------------------------------------------------
     sigma_u |  1.3162176
     sigma_e |  .21703032
         rho |  .97353114   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(2614, 4131) =    14.78           Prob > F = 0.0000
```